Exploring Left-Wing Extremism on the Decentralized Web: An Analysis of Lemmygrad.ml

Utkucan Balci 1 Michael Sirivianos 2 Jeremy Blackburn 1
1 Binghamton University
2 Cyprus University of Technology
ubalci1@binghamton.edu, michael.sirivianos@cut.ac.cy, jblackbu@binghamton.edu

Abstract
This study investigates the presence of left-wing extremism on the Lemmygrad.ml instance of the decentralized social media platform Lemmy, from its launch in 2019 up to a month after the bans of the subreddits r/GenZedong and r/GenZhou. We conduct a temporal analysis on Lemmygrad.ml’s user activity, with also measuring the degree of highly abusive or hateful content. Furthermore, we explore the content of their posts using a transformer-based topic modeling approach. Our findings reveal a substantial increase in user activity and toxicity levels following the migration of these subreddits to Lemmygrad.ml. We also identify posts that support authoritarian regimes, endorse the Russian invasion of Ukraine, and feature anti-Zionist and antisemitic content. Overall, our findings contribute to a more nuanced understanding of political extremism within decentralized social networks and emphasize the necessity of analyzing both ends of the political spectrum in research.

Introduction
In the evolving context of social media, the decentralization of the Web has emerged as a pivotal transformation by creating additional space for online communities that advocate for a more democratic and user-empowered Internet. Although this paradigm shift can promote innovation and freedom of expression in many contexts, like many centralized social media platforms (e.g., Twitter and Reddit), examining how extremists spread their ideology (Pérez Curiel 2020; Benigni, Joseph, and Carey 2017; Zannettou et al. 2018b; Grover and Mark 2019), current work on extremism on these platforms mainly aim to understand the characteristics of far-right extremism online (Scrivens, Davies, and Frank 2020; Zannettou et al. 2018b; Grover and Mark 2019), explore the role of social media in radicalization processes (Ribeiro et al. 2020), and examine echo chamber effects (Efstratiou et al. 2023). Parallel to the discourse on online extremism, the concept of the Decentralized Web has gained attention from researchers, with an emphasis on its potential to democratize social media platforms and mitigate the risks associated with centralized data control (Kwet 2020; Abbing, Diehm, and Warreth 2023; La Cava, Greco, and Tagarelli 2022). Focusing on their peer-to-peer nature and community-driven governance models, research (La Cava, Aiello, and Tagarelli 2023; Struett et al. 2023; Bin Zia et al. 2022; Raman et al. 2019) on Mastodon and the broader Fediverse have highlighted the architectural, behavioral, and operational distinctions that set these platforms apart from their centralized counterparts. While extensive research has been conducted on right-wing extremism within decentralized net-
works (Zannettou et al. 2018a; Gerster et al. 2023; Thiel and McCain 2022), the exploration of far-left extremism on these platforms is still in its infancy.

What is Lemmy? Lemmy is an open-source, self-hosted Decentralized Web platform that provides instances for news aggregation and discussion forums (Lemmy 2023). These instances operate similar to centralized news aggregation platforms like Reddit, where users can create posts, comment on them, and cast votes. Additionally, Lemmy enables its users to communicate across different instances. To date, there is limited literature on Lemmy. (Nunes et al. 2023) analyzed sentiments on Lemmy to understand user migration from Reddit with analyzing nearly 50K posts without touching its extremist content. Notably, this work find positive sentiments toward Lemmy and criticism of Reddit among the users of this platform.

What is Lemmygrad.ml? Lemmygrad.ml is a Marxist/Leninist Lemmy instance that mirrors Reddit and hosts various left-wing online communities. Similar to other Lemmy instances, users in this platform can create communities to post content and engage in discussions around them, by abiding the rules set by Lemmygrad.ml’s administrators. To date, there has not been any research exploring Lemmygrad.ml.

Dataset

Using a custom crawler, we collect 242,370 posts from 1,744 communities (Lemmy’s equivalent of subreddits) between April 20, 2019, and April 30, 2022. Figure 1 displays the top 20 communities of Lemmygrad.ml according to their total number of posts. The two most popular communities, c/genzhourage and c/genzedong, are related to left-wing extremist subreddits that have faced restrictions from Reddit, where r/GenZedong was quarantined for spreading misinformation (Chow 2022), and r/GenZhou was banned for ban evasion related to r/GenZedong’s quarantine (Reddit 2022). Our manual inspection on the banners of these subreddits finds that they fit Petterson’s description of contemporary tankies (Petterson 2020), where they declare their support for Actually Existing Socialist (AES) countries, while also identifying themselves as Marxist-Leninists. The most popular community, c/genzhourage (48,231 posts), is an archive of the subreddit, r/GenZhou. Moreover, the second most active community, c/genzedong, further solidifies Lemmygrad.ml’s role as a Reddit alternative for tankies.

Temporal Analysis

We analyze the temporal information of Lemmygrad.ml, focusing on their popular tankie communities. We first examine monthly user activities to gain an understanding of the Lemmygrad.ml’s growing popularity and the impact of quarantines on tankie subreddits. Subsequently, we explore the evolution of highly toxic posts to complement our findings from the analysis of monthly user activity. Monthly User Activity. To investigate the impact of quarantines on the popularity of r/GenZedong and r/GenZhou on Lemmygrad.ml, we examine the monthly active user counts (Figure 2a) and the the monthly post counts (Figure 2b) for c/genzhourage, c/genzedong, and all of Lemmygrad.ml. We find that c/genzedong shows a sharp rise in monthly user activity after r/GenZedong’s March 2022 quarantine, peaking at 8,637 posts from 1,003 users in April 2022. Before the month of quarantine, c/genzedong’s highest user activity was in January 2022, with a maximum of 96 posts from 58 users. Conversely, as it is an archive community, the ban of r/GenZhou in April 2022 resulted in a heavy decline in c/genzhourage’s user activity. The total number of posts dropped from 7,862 to 41, even though the count of active users peaked, rising from 2 to 19 monthly active users in the same period. We also notice a general uptick in Lemmygrad.ml’s monthly posts, especially post-bans/quarantines, with the numbers rising from 16,944 posts by 1,130 users in February 2022 to 24,282 posts by 2,529 users in April 2022. Our findings suggest that, similar to their right-wing counterparts on Gab (Thiel and McCain 2022), platform restrictions greatly increased the popularity of Lemmygrad.ml.

Monthly Perspective Analysis. Perspective API (Perspective API 2018) is a tool developed by Google which measures abusive or hateful content within texts, providing scores from 0 to 1, lowest to highest. Using Perspective API, previous research on Mastodon (He et al. 2023) and Pleroma (Hassan et al. 2021) found the existence of highly toxic posts within these networks. Ribeiro et al. (Horta Ribeiro et al. 2021) observed an increase in toxicity levels following the migration of users from the subreddit r/TheDonald to the platform thedonald.win, using Perspective API scores as a metric for measurement. Building on this, we measure the monthly high Perspective API score proportions of the posts on Lemmygrad.ml. Following previous work (He et al. 2023; Hoseini et al. 2023; Balci et al. 2023), we consider posts $\geq 0.8$ as having high Perspective API scores. We use the production models (Jigsaw 2024) of

![Figure 1: Top 20 communities of Lemmygrad.ml. Communities related to tankie subreddits are colored as red.](image-url)
Perspective API, which have undergone testing processes in various domains and have been developed based on large volumes of comments annotated by humans.

Figure 3 illustrates the monthly proportions of high scores from Perspective API models for c/genzhouarchive, c/genzedong, and all communities of Lemmygrad.ml. We find that in 2019, the year Lemmygrad.ml was founded, there was a notable peak in the frequency of high scores across TOXICITY, SEVERE_TOXICITY, INSULT, IDENTITY_ATTACK, THREAT, and PROFANITY models. This peak occurred despite 2019’s relatively low monthly activity, which averaged 365 posts. However, the subsequent years, 2020 and 2021, saw a substantial reduction in the average high scores for all models, cutting them by more than half.

Interestingly, in March and April 2022, despite recording the highest number of monthly posts, the mean high scores for all perspective models peaked for the first time since 2020. This increase suggests that the platform migration of tankies led to an increase in toxicity on Lemmygrad.ml.

**Takeaways.** We find left-wing extremists have substantial activity on Lemmygrad.ml. Our analysis shows that Lemmygrad.ml’s top two communities, c/genzedong and c/genzhouarchive, are associated with tankie subreddits, r/GenZe-dong and r/GenZhou. Furthermore, our findings suggest that tankies’ platform migration resulted in an increase in user activity and toxicity on Lemmygrad.ml, highlighting the importance of considering the impact of platform migration to decentralized networks on online communities.

**Topic Analysis**

We aim to understand the content discussed in this platform by using a transformer-based topic modeling technique and explore their popular discussions to gain an understanding of Lemmygrad.ml.

**Topic model.** We leverage BERTopic (Grootendorst 2022), which uses transformer-based sentence embeddings to generate embeddings of documents and has been previously used in studies to explore content of online communities (Hanley, Kumar, and Durumeric 2023; Mekacher, Falkenberg, and Baronchelli 2023; Tahmasbi et al. 2024). This method automatically finds dense clusters of documents by first applying dimensionality reduction to the embeddings with UMAP (McInnes, Healy, and Melville 2018), and clustering them using HDBSCAN (McInnes et al. 2017). Before training our model, we remove hyperlinks and user names from posts, and exclude deleted or removed posts. Additionally, we observe that posts from the c/genzhouarchive contain the text: “originally from r/Gen-Zhou,” so we also remove these references from the posts. To refine our analysis, we also manually inspect 100 random posts from each topic.

**Popular discussion themes of Lemmygrad.ml.** Table 1 presents the top 20 topics with most posts on Lemmygrad.ml. We find that users on Lemmygrad.ml frequently discuss two AES countries, China and North Korea, with many posts expressing support for them. An example post from c/genzedong:
In the topic related to Ukraine, we find many posts supporting or justifying the Russian invasion of Ukraine. An example post from c/politics:

“No one’s asking you to be ok with all the horrors of war. But if we don’t understand what started it, we won’t be able to finish it. Ukraine’s ethnic cleansing of Russians, suppression of the LPR and DPR, and flirtation with NATO must stop if ever this war will end.”

We find that discussions on the Israel-Palestine conflict primarily criticize Israel. While many posts condemn antisemitism, we also encounter numerous posts that extend beyond criticizing Israel, displaying anti-Zionism and even antisemitism. An example post from c/communism:

“1. Jews aren’t a nation. 2. Israel is a fake country. 3. The only thing which could be more fake than Israel and the “Jewish” national identity would be for the fans of start trek in to create a language of their own similar to English, then enter the bourgeoisie and amass wealth and power, and then invade some imperialized nation 200 years later, settle in it and claim their own stark-treakish fake nation state. And they will justifying this by quoting some science fiction book written in the 1900s which tells how are they the chosen people to colonize this new land. Everyone identifying as a jew is a zionist and an enemy of the communist and anti imperialist movement.”

Besides relatively mundane topics and those related to hobby-related communities, we see discussions related to cryptocurrencies and vaccines. We also encounter topics related to ideologies, economics, and sex work, where posts predominantly reflect Marxist/Leninist or anarchist perspectives.

**Takeaways.** Our results show that users of Lemmygrad.ml frequently share posts that support authoritarian regimes, as seen in their support for China, North Korea, and Russia. Moreover, their support can extend beyond backing these authoritarian regimes, even cheering on their violent actions, as evidenced by their posts on the Russian invasion of Ukraine. Additionally, we observe anti-Zionist and antisemitic behaviors, which show similarities to right-wing extremism (Zannettou et al. 2020). Our analysis suggests a concerning endorsement of authoritarian actions and extremist rhetoric on Lemmygrad.ml, further indicating that left-wing extremist communities on decentralized platforms should receive more attention from the academic community.

**Conclusion**

In this paper we explored a left-wing Lemmy instance, Lemmygrad.ml, which also serves as a hub for left-wing extremist subreddits that faced restrictions from Reddit. We find an increase in user activity and toxicity levels on Lemmygrad.ml following the migration of r/GenZedong and r/GenZhou. Furthermore, our analysis of the content revealed posts supporting authoritarian regimes, endorsing the Russian invasion of Ukraine, and exhibiting anti-Zionist and antisemitic rhetoric. Our findings underscore the importance of studying left-wing extremism on decentralized platforms alongside right-wing extremism to gain a comprehensive understanding of the full spectrum of political extremism on the Decentralized Web. Additionally, by investigating the decentralized nature of platforms like Lemmygrad.ml, this paper contributes to the broader discourse on the implications of the Decentralized Web technologies.

**Future Work**

While this work highlights the presence of left-wing extremism on the Decentralized Web, it is limited to a single instance on a single platform. Our study emphasizes the critical need to explore the prevalence and dynamics of left-wing extremism within the Decentralized Web. We also encourage future work to further investigate the distinctions between left-wing and right-wing extremism on these platforms, as well as to compare the characteristics of left-wing extremists on decentralized platforms with those on centralized social media platforms.

**Acknowledgments**

This material is based upon work supported by the National Science Foundation under Grant No. IIS-2046590 and the European Union under the project MedDMO (Grant Agreement no. 101083756).

**References**


